Redesigning the Census Long Form: Results from the 1990 Alternative Questionnaire Experiment

Theresa J. DeMaio and Nancy A. Bates
U. S. Bureau of the Census

In any survey, the data collection instrument is key to the quality of information that is collected. The census form is no exception. Since the census is conducted largely through mailout/mailback methods involving a self-enumerative questionnaire, the form itself largely determines how well people respond to the census--whether or not they participate, and the completeness and accuracy of the data they provide.

Since 1985, the Census Bureau has been engaged in a multi-stage program of research aimed at improving the data collected in the census, specifically in terms of improving mail response rates, item nonresponse rates, and data quality. This research is focused specifically on the census long form, and has included a number of different techniques to obtain information about respondent problems with the form and to test potential solutions to those problems. Initially, one-on-one interviews were conducted, with observers carefully watching subjects as they completed the form and then debriefing them afterwards (see DeMaio, 1986, for a discussion of the results of this research). This was followed by a series of small split-panel experiments in which a revised version of the form was tested against the original form (see DeMaio, Martin, and Sigman, 1987, and Martin, DeMaio, and Campanelli, 1990 for further information). The sample for these experiments was relatively small (about 500), not statistically representative, and used a different mode of data collection than that used in the census itself. Therefore, the next step in the research was to conduct a large nationally-

representative field test using a mailout/mailback methodology (see Bates and DeMaio, 1989, and Bates and DeMaio, 1992 for descriptions of this test).

At each of these steps, the scope of the research broadened slightly. From using a single pre-existing census form, the mailout/mailback test involved three alternative forms and a control, so that changes in layout, question wording, and question order could each be assessed separately. Even with that test, however, a critical aspect of the "census" experience was missing--the publicity campaign that surrounds the taking of an actual census. To remedy that shortcoming, we conducted a split-panel experiment within the context of the census itself. The results of that experiment--called the 1990 Alternative Questionnaire Experiment--are presented in this paper. After describing the questionnaires that were included and the methodology for the test, we present results relevant to the three main objectives of the research project: improving mail response rates, item nonresponse rates, and data quality (specifically as it pertains to the coverage questions).

QUESTIONNAIRES

Six questionnaires were included in the Alternative Questionnaire Experiment. Along with a control form, five experimental questionnaires were developed, which included varying levels of departure from the design of the 1990 census form.

The control form (Panel 1) was identical to the 1990 long form. It was designed to be a FOSDIC look-alike form, however, and the data were not captured through the automated processing equipment. It had a fold-out flap on which a listing of all household members was obtained, followed by a matrix of short-form information collected about every person in the

household, then three pages of housing information (both 100-percent and sample questions), and finally two pages of sample population questions for each household member.

The form for Panel 2 followed the same basic construction as the control form, but it introduced some minor changes designed to make the form more "respondent-friendly." There were three types of changes: 1) step-by-step instructions printed in red ink to help guide respondents through the various sections of the form; 2) minor wording changes to simplify and clarify concepts; and 3) graphic changes to increase the amount of white space on a page.

The form for Panel 3 was the same as Panel 2 with two exceptions: two question sequence experiments were embedded within the 100-percent person matrix of the Panel 3 questionnaire. First, the relationship question was reordered from first to third, following the sex and marital status items. This resequencing was performed to improve item response to the relationship item by first giving respondents a chance to answer the simpler and more straight-forward sex and marital status questions. Second, the order of the race and Hispanic origin questions was reversed. This was done because Hispanic origin typically has the highest item nonresponse rate in the 100-percent person section, and anecdotal evidence as well as previous research (see Martin, DeMaio, and Campanelli, 1990) suggests that respondents feel this question is redundant once they have answered the race question.

Beginning with Panel 4, more radical changes were introduced into the design of the experimental forms. In Panel 4, the basic structure of the form was abandoned, as well as the fold-out flap. The basic purpose of this form was to keep all the person items and housing items together, and thus simplify the respondent's path through the form. This was done by creating a booklet in which all the person questions were arranged in a matrix format. Following the

household roster, names of household members were entered once across the top of the booklet. Then, respondents could turn shorter pages that were aligned with columns for each of the persons listed above, and which contained response spaces for each of the person questions. Following the person section, housing questions were placed at the back of the form. The minor wording and format changes introduced in Panel 2 were incorporated in Panel 4 also.

In Panel 5, the concept of a single form was abandoned. In this panel, respondents received a kit consisting of a folder and a number of separate forms. The purpose of this design was to encourage self-response, so it would be easier to obtain responses from households containing unrelated members, or even related household members for whom privacy was an issue. In Panel 5, the instructions for completing the forms were printed on the folder, and nine person forms and one housing form were included inside. The individual person forms contained both the 100-percent and sample questions in a small fold-out form like an Individual Census Report. Information about two more people could be accommodated in this way, since the traditional census format only has space for seven persons. The housing form contained four regular-sized (8 1/2 x 11) pages of questions, including the household roster, 100-percent housing items, and sample housing items. In addition to changes in the basic structure of the form, revisions were made in the household roster and coverage questions. The focus of these questions was changed, to elicit reporting of marginal and/or temporary household members. These changes will be discussed in detail later in the paper. As for the other questions on the form, the minor wording and format changes introduced in Panel 2 were incorporated in Panel 5 also.

The final questionnaire, Panel 6, used the same basic kit-style format as Panel 5, and

incorporated the minor question wording and format changes maintained in the other experimental panels. The only difference between Panels 5 and 6 was that the latter did not request names or other identifying information. The reason for the design of this form was to alleviate confidentiality concerns, which have consistently surfaced as a source of threat to respondents in focus group research conducted by the Census Bureau. In addition to the household roster, the relationship and place of work items were also omitted from this form, so it is completely anonymous.

METHODOLOGY

The Alternative Questionnaire Experiment was conducted among a representative sample of all households in densely populated, central city areas of the country. Approximately 7,000 households were randomly selected to receive each questionnaire version, for a total of 42,000 households participating in the experiment. Each household received an experimental form instead of the actual census form.

Census forms for sample households were mailed from the Census Bureau's clerical and data processing facility in Jeffersonville, Indiana. (In contrast, regular census forms were sent directly by the printer to the appropriate local post office.) The AQE forms were returned to the Jeffersonville facility as well, while most census forms were sent to one of seven large processing offices located across the country. (Forms were returned to the office nearest to the address.) These procedural differences between the experiment and the census may have affected the results, as will be discussed below.

Once the forms were returned to the Census Bureau, they were checked in, coded, and

keyed. To evaluate how well the forms were completed by respondents, we looked at unedited responses without any imputation. Some additional coding was done to capture information about how well respondents completed certain portions of the form. Statistical analysis was conducted using SAS, assuming that data were collected from a simple random sample. The sample for the AQE was selected using a systematic random sample, but information about persons in the household is clustered. SAS does not adjust for this clustering effect, and to the extent that persons within a household have similar characteristics, the sampling errors estimated under the assumption of simple random sampling will underestimate the true sampling errors for a clustered design. Therefore, all significant differences found among the person data were subjected to more stringent analysis using VPLX, variance estimation software for complex samples.

RESULTS

Mail Response Rates

As noted previously, questionnaires for the AQE were mailed out from and back to different locations than the rest of the census questionnaires in the country. As a result, two logistical problems ensued. First, there was a greater lag time for AQE households than for others between time of mailing and time of delivery both for blank questionnaires mailed out and completed questionnaires being mailed back. Forms took longer to get to respondents, and they also took longer to get back to be processed. Second (and more troublesome), because of the large volume of mail being returned to the census processing offices and because the AQE envelopes were generally similar to the regular census envelopes, there was a potential for the

experimental forms to be mistaken for regular census forms and returned to the wrong location.

Due to these two factors, the mail response rate for the households in the AQE lagged behind the corresponding level of returns for the nation as a whole. We suspect, but were never able to prove, that AQE questionnaires were delivered to the incorrect census processing offices, and that they got lost in the shuffle of the huge volume of mail. By the time nonresponse follow-up began on April 25, the national mail response rate for long forms from central city offices was 53 percent, and the average mail response rate for the experiment was 47 percent. However, despite these differences, for the most part these logistical problems affected all of the forms in the experiment; therefore, the comparison between panels was not compromised. We continued to receive completed questionnaires after the April 25 cutoff date for field followup in the census, and we accepted them in the database. Thus, the final tally of mail response rates was 51 percent.

Table 1 shows the rates of mail response for the various alternative questionnaires. As can be seen in this table, the control panel had the lowest mail response rate (48.2 percent), and the level of response increased as the magnitude of the changes to the form increased.

Only slight differences were observed among the first three panels, and these differences were not significant. The forms had the same basic structure, and the revisions that were made in Panels 2 and 3 were relatively small (addition of step-by-step instructions in red ink, changes in question wording, and changes in the order of questions). These changes may have made a difference to respondents who were already disposed to respond to the census, but in the final analysis, the structure itself was what mattered.

The format of Panel 4 was the first real departure from the traditional structure. Here

also, the first notable increase in response occurred. The mail response rate was 51.8 percent, significantly higher than the control form. In Panel 4, the complex folding pattern was eliminated by removing the flap, and the first two inside pages were simple, uncluttered pages containing instructions and the household roster question. The form was basically a booklet with a matrix of population questions for the first 17 pages, with two pages of housing items following that. Thus, the higher response rate would seem to be a function of these changes. However, other explanations are possible. One implication of the booklet design was that the form was mailed out flat rather than folded, and as a result a larger envelope was used. The form was mailed back flat as well. The larger envelope itself could have affected the mail response rate, either because it increased the likelihood that respondents would notice the form in their mail, or because it differentiated the forms in this panel from the volumes of mail going to the processing office. We cannot determine the extent to which either of these potential explanations occurred; however, the Panels 5 and 6 return envelopes were the same size as the control envelopes and as we shall see they also had higher return rates, so visibility by the post office is not likely to be sufficient cause for the increased rate. Thus, to the extent that the larger questionnaire size was necessary to accommodate the booklet design, these features worked together to increase mail response.

Panels 5 and 6 contained a different kind of departure from the standard format--a "kit" approach. As Table 1 shows, this approach was also successful in increasing mail response. While the small increase in response between Panel 4 and Panel 5 response rates was not significant, the rate of response for Panel 5 (52.5 percent) was significantly higher than for the control. This suggests that the individual forms approach, which encourages self-response and

eliminates the problem of trying to follow a road map through a long, complicated household form, has potential as a design for future censuses. In Panel 6, the kit approach was also used and again, response to this form was better than to the control. In addition, the nameless aspect of the form seems to have stimulated additional response. The response rate for Panel 6 as almost two percentage points higher than for Panel 5, a difference which is statistically significant; however, the effect of providing anonymity to respondents is smaller than the increase in mail return due to making structural changes in the questionnaire. This effect of providing anonymity is smaller than we might have expected, but it still lends support to the findings from focus group research conducted by the Census Bureau over the years that indicate concerns about confidentiality. And it suggests that further research should focus on issues in this area.

Item Nonresponse

Nonresponse to Population and Housing Sections

Mail response is not the only indicator of the level of completeness of census data. Respondents who complete the form do not necessarily complete the entire form--they may leave either individual items or whole sections blank. And some designs may lend themselves to higher levels of completeness than others.

We examined completeness of response in two ways. First, we broke the form up into sections (100-percent person, sample person, housing) and calculated the percentage of cases in which all the items in the section were left blank. Second, we looked at individual items to see how completely they were answered.

Table 2 contains the results of our analysis of nonresponse by sections of the questionnaire. Nonresponse for the 100-percent and sample person sections together is presented in the second row, and the percent of cases that were missing 100-percent data only and sample data only are shown in the next two rows. Review of these three rows shows that the pattern of missing sections of data varied dramatically with the basic structure of the form. Panel 1 had very few cases in which either all the population items or only the 100-percent items were blank. However, a relatively large percentage of cases (4.3 percent) were missing sample person data. This general pattern is repeated for Panels 2 and 3; however, for these forms the differences were not so extreme. Significantly higher percentages of the forms are missing either all population items or just the 100-percent items compared to the control, and slightly (but not significantly) fewer were missing the sample population items.

The pattern for Panel 4 is quite different. Here, although the percentage of forms missing all the population data was more than three times
higher than the control (1.8 percent vs. 0.5 percent), the levels of other kinds of missing data compared favorably. Very few returned forms were missing all the 100-percent population items, and the percent missing all sample data was the same as the percent who missed both the 100-percent and sample data (1.8 percent). While this percentage was high compared to the control in the latter case, it was less than half of the corresponding percentage for the control in the former case. This pattern makes sense given the design of the form--once respondents began to complete the person questions, there was no obvious differentiation between the 100-percent and sample sections.

In Panels 5 and 6 which use the kit format, another pattern of missing data by sections is

evident. For both of these forms, the incidence of noncompletion of all the population items (both 100-percent and sample) was high. The level of missing data was particularly high for Panel 5 (5.2 percent), and it is not immediately clear why the omission of identifiers in Panel 6 should decrease the level of missing data by more than half (to 2.2 percent). However, it is not surprising that in both these panels the majority of nonresponse to the population items would be due to complete nonresponse (to both 100-percent and sample sections) rather than to either one of these sections alone. The individual person form contained both sections, and the main deterrent to response seems to be losing the form or for some other reason mailing back a kit that was missing some person forms. The form itself was straightforward with no skip patterns, and it was immediately clear exactly how many (or how few) questions were asked about each person. As Table 2 shows, nonresponse to either of the individual sections was rare.

Levels of nonresponse to the housing data also differed by questionnaire version. Nonresponse to all the housing items was relatively low for the first three panels, which used the traditional structure. However, in each of the alternative structures, the rate of nonresponse to all the housing items significantly increased. This is not too surprising given the design of the forms.

The Panel 4 form contained all the housing questions at the back, where they could have been overlooked after the population items were completed. The results in Table 2 suggest that this did indeed happen, since the nonresponse rate of 4.0 percent for this panel was more than double the rate for the control panel.

In Panels 5 and 6, the housing questions were contained on a separate housing form. With this structure, there was a possibility that the housing form could be lost or misplaced,

resulting in high levels of missing data for the entire section. As with Panel 4, this structure did result in higher rates of missing data for the entire housing section. It is interesting to note that Panels 5 and 6 had generally opposite patterns of entire section nonresponse for the population and housing sections. For Panel 5, 5.2 percent of population forms and 3.1 percent of housing forms were not returned. For Panel 6, a higher percentage (5.3 percent) of housing forms did not make it back to the processing office, while only 2.2 percent of the population forms were missing from the returned kits. It is understandable that the elimination of identifiers would increase the likelihood that population forms would be returned, but it is not clear why it would have an effect of decreasing the completion of housing information.

It is clear by looking at the mail response levels and the levels of nonresponse to entire sections of the form that there are tradeoffs among the forms in these two aspects of nonresponse. The next question is, what is the bottom line in terms of total nonresponse, taking both of these aspects into consideration. The last three rows of Table 2 show that for the population items, the gains in form mail response for Panels 4 and 6 more than compensated for the losses in entire section response. For these two panels, the comparison for both 100-percent and sample population items was significantly favorable compared to the control. While other of the panels (i.e., Panel 3 and Panel 5) showed some positive differences, in neither case were both differences significant.

As far as the housing items are concerned, differences between the forms in total nonresponse were less pronounced, although still statistically significant. Here Panel 5 fared the best; Panel 6 and Panel 3 also had significantly less nonresponse than the control.

Nonresponse to Individual Items

The lowest level of item nonresponse involves omitting a response to an individual item. To conduct this analysis, we calculated item nonresponse rates for individual items, excluding cases where the entire section that contained the item was left blank. For purposes of presentation here, we calculated summary statistics to measure average item nonresponse.[1] Summary statistics for the 100-percent person section, sample person section, and housing section are presented in Table 3. The summary statistics are sensitive to the fact that people are required to answer a different number of items depending on their situation (e.g., owners vs. renters in the housing section; age and labor force characteristics in the sample person section).

The first row of Table 3 shows that there were differences overall in the level of nonresponse to the 100-percent person section. Nonresponse on the control panel was the highest, with 5.8 percent of the 7 person items left blank, on average.[2] Nonresponse generally declined with each succeeding questionnaire panel--for Panels 5 and 6, nonresponse was reduced by almost two thirds.

A number of question wording, layout and sequence changes in Panels 2 and 3 were responsible for the decrease in average nonresponse. The extremely high rate of missing data for Panel 1 is largely due to the fact that 19 percent of the responses for Hispanic origin were missing. In Panel 2, the average nonresponse rate declined by more than half with the addition of an instruction to "Fill in the NO circle if not Spanish/Hispanic." Additionally, in Panel 3 it declined further with the reversal of the race and Spanish origin items (see Bates, 1991, and Bates 1992, for further discussion of the individual item nonresponse rates).

Other changes such as the sequence experiment that moved the relationship item down to

third position in Panel 3 also had a positive effect. Conversely, the alternative format of the age and year of birth items did not have positive results.

The wording and sequencing of 100-percent items in Panels 4, 5, and 6 were unchanged, and similar to Panel 2. However, significant decreases in average item nonresponse were observed. This suggests that the general layout affected the level of missing data over and above the effects of changes to particular questions or their order. The booklet format of Panel 4 and particularly the individual-person-forms approach of Panels 5 and 6 were successful in simplifying the response task and increasing the amount of information reported.

The middle row of Table 3 suggests that the positive effects of the individual-forms-approach carry over to the sample person section, since the average levels of item nonresponse for Panels 5 and 6 are significantly lower than for all other panels. This is not surprising, since this format eliminates all the complexities of having to find one's way through the form, and at the same time it clearly indicates how many questions have to be completed for each person.

The design of Panel 4, however, did not show any improvement over the control in terms of the average percent item nonresponse to the sample person section. Although all the person items were together, simplifying the path through the questionnaire, the advantage of this format did not seem to hold the interest of respondents. Perhaps the length of the form overwhelmed respondents by the time they got past the 100-percent person section.

The final section of Table 3 presents average item nonresponse rates for the housing section. The first row suggests that there was a significant difference over all the panels, with Panel 4 being the outlier and all the other panels being relatively equal. This seemed somewhat

odd to us, as we could not hypothesize why this questionnaire might have a higher average nonresponse rate once the whole-section nonrespondents had been removed. However, a closer examination of the item-by-item nonresponse rates revealed a plausible explanation for the difference.

In Panel 4, nonresponse to each item in a series of four questions about utility costs was noticeably higher than other panels. This was not characteristic of other housing items in this panel. A typographical error was found in this series which could explain the higher-than-usual nonresponse rates. When we removed these items from the analysis, the results in the bottom line of the table were observed. The difference between Panel 4 and all others dropped dramatically and the nonresponse rate for Panel 4 did not differ from that of the control. Thus, moving the housing questions to the back increased the frequency of whole section nonresponse but did not appear to increase average item nonresponse among respondents who answer the section.

The results for Panels 5 and 6 excluding utility items indicate a slight decrease in nonresponse to the housing items. Although the overall Chi-square was significant, the differences were relatively small. In general, individual item nonresponse to items in the housing section seems unaffected by the design of the form.

Data Quality

The third area in which we looked for improvement was data quality. In this paper, we examine data quality as it relates to the coverage items.

The major purpose of the census, of course, is to count the nation's population. In that

regard, the most important piece of information collected is the household roster, which determines how many people get counted. Historically, the census has undercounted certain population subgroups, in particular young Black males. The content of the census form is not generally perceived as being responsible for the undercount; nevertheless, in designing the alternative census forms we made some changes that were focused on aspects of the form related to coverage.

Traditionally, the household roster is the first item on the census form, preceding the 100-percent population items. The changes we made to this item varied across the different panels. In the control form (Panel 1), the household roster item was located on the foldout flap. The question was written is blue ink and followed by lists of residency rules written in black ink. The household roster itself was formatted in a double-column style. In Panels 2 and 3, the revisions included graphic and format changes only. First, a step instruction printed in red was placed before the question, telling respondents to "Make a list of the people who live here." Second, the question itself and the residency rules were shaded in light blue, to be consistent with the design of the rest of the form. Finally, a single column format was used to list the names.

In the booklet design of Panel 4, the flap was eliminated. The household roster question was on page 3 of the booklet, facing a page of short instructions about how to get started. The step instruction, light blue shading, and single column format were also maintained. For Panels 5 and 6, the household roster was included on the housing form rather than the person form, since it is asked only once per household. In Panel 5, the residency rules were omitted, and names of household members were requested in a single-column format. In Panel 6 (the

nameless panel), no roster was included at all. Instead, a question was added that asked about the number of people living or staying at the household.

Changes were also made to the coverage questions that traditionally come at the beginning of the housing section, after the 100-percent person matrix has been completed. (Attachment 1 contains examples of all the coverage items used in the experiment.) The coverage question on the control form (Panel 1) consisted of two parts: one asked about persons left off the form who should have been counted, the other asked about people who were included on the form but perhaps should not have been.

In Panels 2 and 3, these two parts were separated, and the overcoverage and undercoverage aspects of the concept were dealt with in different ways. At the beginning of the housing section, a question that asked about persons left off the household roster was included. Then, in the sample person section, a question was added that was meant to probe for potential overcoverage. The question asked whether there was another place where the person lived all or most of the time, and the response categories reflected the content of the residence rules. That is, they were designed to identify persons who may have been included on the household roster when they should not have been, because they were away at college or boarding school, at a military base, etc.

In Panel 4, the same strategy was employed, dealing separately with undercoverage and overcoverage. The difference was that in this panel, two questions were added in the sample person section: one asked for the person's living situation as of April 1 (i.e., lives here all or most of the time, lives somewhere else most of the week while working, lives here some of the time, visiting or staying here temporarily), and the second was the same question included in

Panels 2 and 3.

In Panel 5, the strategy was expanded somewhat. No residence rules were included in the household roster question. And the question getting at undercoverage was revised to focus explicitly on persons with marginal attachments to the household, who might not have a permanent attachment anywhere and thus might not get counted in the census. On the overcoverage side, a third question was added to the questions in the sample person section. This question asked for the address of the other place where the person lived most of the time. The purpose of this was to facilitate checking on duplicate counting of these persons in the census. (This has not been done, although it might be done in the future.)

In Panel 6, our attempt to manipulate the coverage questions was extremely limited, since no names or other identifiers were obtained. The undercoverage question was eliminated, although the instructions did include, on both the population and housing forms, a reminder to complete forms for both regular and marginal household members. The two overcoverage questions included in Panel 4 were also included on the population form.

Table 4 presents the item nonresponse rates for the household roster item. It is evident that both the format changes in Panels 2, 3, and 4 and the switch to the individual forms in Panels 5 and 6 resulted in improvements in respondents' likelihood of completing the item. The item nonresponse rates for Panels 2, 3, and 4 were similar, and each was less than half of the rate for the control panel. This suggests that the graphic changes were successful, and that the removal of the flap was less important in terms of completing the roster item. The level of missing data for the individual forms was extremely low, with less than one percent of the forms having no response to the roster. (The figure for Panel 6 represents level of response to the item

requesting the number of household members, since a roster of names was not included on this panel.)

Just as important as the level of missing data for this item is information about the persons counted at the address. Table 4 shows that overall, there was a statistically significant difference in the mean number of household members reported.[3] However, this difference was extremely small. In terms of individual comparisons, the only significant improvement over the control occurred in Panel 5. The format of this panel omitted the residency rules from the original question, but included a second roster designed to elicit names of marginal residents. These results suggest that this format is successful in stimulating increased reporting of household members.

The other manipulation that we attempted in this experiment was also marginally successful. We included nine person forms rather than seven in the kits for Panels 5 and 6, and we hypothesized that this might encourage more reporting. As noted above, the average household size reported was significantly larger than the control for Panel 5 but not for Panel 6. However, we also looked at the percentage of cases in which eight or more names were included on the household roster. This comparison is of interest, since person information is obtained for only seven persons using the traditional structure of the census form. (Information about additional persons in households with more than seven persons is obtained separately.) Here, we see that the overall comparison of the number of cases with more than seven persons was statistically significant, although the differences across experimental panels were relatively small. Both Panels 5 and 6 had higher rates of reporting more than seven persons than the other panels. Thus, these results present some evidence that lengthening the form to accommodate

more people may be associated with higher reporting on the household roster. We also explored the possibility that dropping the name requirement would stimulate respondents to more frequently report males, who are more frequently undercounted. To do this, we examined sex ratios for the 15-29 age group, and we were looking for a higher ratio of men to women in Panel 6 as evidence that our hypothesis was supported. However, as Table 4 shows, this was not the case; the sex ratios did not differ across the panels.

Thus, we have some evidence that our experimental manipulations within the household roster were successful, but the nameless panel was not overly effective in stimulating increased reporting of persons.

In terms of the coverage questions, we first examine the results of the experimentation relating to undercoverage. Table 5 presents the percentage of forms in which respondents reported that they left names off the household roster because they were not sure whether the persons should be listed. Overall, there were significant and sizable differences in the extent of potential undercoverage across forms. The control panel demonstrates the lowest level, with all the question wording and placement variations increasing the percent of forms with names reported. This suggests that, following standard census procedures to follow up on any form that has additional persons reported, the alternative forms could potentially result in fewer missed persons.

The revised wording of the undercoverage question in Panels 2, 3, and 4 showed a slight but significant improvement over the control in increasing the reporting of possible undercoverage cases. The major gains in reporting, however, were elicited by the expanded coverage question in Panel 5, which was expanded to focus on persons who might get missed in

the census. In this panel, the placement was also different, since it was at the bottom of the separate housing form, in a more prominent location. Thus, either one of these factors could explain the higher rate of response.

The next set of analyses concerns the questions designed to measure overcoverage in the census. The bottom half of Table 5 presents these results.[4] Again, we see that Panel 1 showed the lowest percentage of potentially "incorrect" reporting in the census in terms of counting people who should have been omitted from the roster. Percentages for all the experimental panels, for which the overcoverage questions were included in the sample person section, were dramatically higher than for the control panel, which contained a single question in the housing section.[5] Panels 2 and 3 had the highest rates of potential overcoverage, with Panels 4, 5, and 6 falling in between. The content of the series of items in the sample person section is likely responsible for some of the differences. Panels 2 and 3 contained a single question that asked "Is there another place this person lives all or most of the time?", while Panels 4, 5, and 6 had contained a contingency screener which first asked "Which of the following best describes this person's living situation as of April 1?" Persons who were reported as living "here all or most of the time" were skipped out of the second question, which is the basis for the figures in Table 5.

Persons in Panels 4, 5, and 6 who answered the second question when they should not have could be identified based on their response to the screener and excluded from subsequent analysis. However, without the screener (in Panels 2 and 3), it was not possible to isolate persons who actually live at another address all or most of the time from those who misunderstood or skimmed the question and answered it incorrectly. To measure the effect of the screener, the percentages of persons "overcovered" or potentially ineligible were calculated

for Panels 4, 5, and 6 first using the screener and then again without it. When the screener was not used to identify the proper base for the "other place" question, the percentage of persons who reportedly live at another house or apartment, college, an institution or a military base increased from 1.5, 2.9, and 3.1 percent in Panels 4, 5, and 6 to 2.6, 4.0, and 4.4 percent, respectively. These latter percentages are more in line with those found in Panels 2 and 3, which did not contain the screener. This suggests that both the screener and the new "other place" item are necessary to determine whether persons fit the Census Bureau definition of "eligible."

The results of these coverage experiments suggest that something very different happens in the alternative method of evaluating overcoverage as compared with the control. On the traditional form, respondents are expected to read the examples, mentally review everyone listed and rethink whether they have included someone who really does not meet the residence criteria. Asked in this way, very few reports of overcoverage are elicited. The alternative forms take a different approach, finding out about each person one at a time by embedding a question about residency patterns within the population questions. Apparently, this is easier for respondents because they are provided with a structured question, complete with response categories for each person--all they have to do is select the one that best applies. The results from asking the question this way indicate a different conclusion--that ineligible inclusions may not be so rare, especially among those who are perceived as having a second house or apartment and for college students living away from home.

The differences between these two methods should not be taken too literally, however, because there are some limitations to this experiment. First, no attempt was made to recontact households and confirm the living situations of those reported in "ineligible" categories. Second,

we do not know for sure whether the sample for the control panel actually contained fewer ineligible persons, or whether the question used in the control panel was simply less sensitive than the alternative question. Finally, since the two panels without the screener question reported a much higher amount of potential overcoverage than the panels with the screener, we must speculate that some respondents either misread or misinterpreted the screener question. "All" or "most" may be ambiguous for people who live/stay half the time in two different places. Thus, the results suggest that some additional work in this area is needed to develop and refine questions to measure and classify residency.

CONCLUSION

In this experiment we devised several alternative strategies to improve response in the census. First, we made small changes in traditionally structured census long forms to make them more respondent-friendly (Panels 2 and 3). Second, we tried two different ways of making dramatic changes in the structure of the form (Panels 4 and 5). And third, we provided anonymity to respondents and their household members (Panel 6).

The results showed that we were successful in some but not all of our revisions. Our attempts to motivate respondents to complete the traditionally-structured form were generally not successful. Mail response rates to Panels 2 and 3 were not significantly higher than for the control (Panel 1). Average item nonresponse to the 100-percent population items was improved by the changes we made, but there was no difference in nonresponse to either the sample population items or the housing items. These results suggest that the structure of the form is a major impediment to respondent cooperation.

The larger changes we made to the structure of the form were much more successful. Both the booklet form (eliminating the flap and placing all the person items together) and the kit approach (individual forms for each person in the household) produced significant improvements in mail response as well as improvements in response to individual items. To be sure, there were tradeoffs between whether or not forms were returned and how completely whole sections were filled out. However, in terms of total nonresponse--taking both these kinds of missing data into account--the booklet format of Panel 4 and the individual-forms approach of Panel 5 demonstrated that dramatically different ways of approaching the design of the census form have potential for increasing respondent cooperation.

Our effort to provide anonymity to respondents also produced significant improvements in mail response. These results demonstrate that, as we hypothesized, confidentiality concerns do affect the amount of data collected in the census. Although a nameless census form is not a feasible approach, the results suggest that future research should concentrate on the issues of privacy and confidentiality, in order to develop other methods of increasing response to the census.

Our experimentation with the coverage items was inconclusive, but it certainly pointed out that a range of issues must be dealt with, especially as far as overcoverage is concerned. This effort was really the first step in dealing with the topic of residency, and the results point out the importance of additional research to develop and refine questions to measure and classify residency.

These results provide a solid foundation with which to begin research to tackle to complicated prospect of taking a census in the Year 2000.

NOTES

[1] The denominator of the statistic is defined as the number of items a respondent should have answered in a section; the numerator is a count of the items left unanswered. By multiplying by 100, we arrived at an average percent nonresponse for the section.

[2] For Panels 1-5 the items were relationship, sex, race, age, year of birth, marital status, and Hispanic origin, with Person 1 in each household being required to answer all except the relationship item. In Panel 6 there were six items, with relationship omitted.

[3] In panels 1-4, the mean number of persons reported on the roster was calculated; in Panel 5, the mean was calculated from the sum of the names in questions 1 and 2 so that household members added as a result of the expanded coverage question could be taken into account; in Panel 6, the response to the question asking about the number of people in the household was used.

[4] Data for Panel 1 come from the overcoverage question in the housing section; figures for the other panels were generated from the new question added to the sample population section.

[5] In the sample person questions, cases in which the person has another address at college, in an institution, at a military base, or at another house or apartment are considered to be potential overcoverage.

REFERENCES

Bates, Nancy, 1990 PREM #146: Additional Results from the 1990 Alternative Questionnaire Experiment, May 7, 1992.

Bates, Nancy, 1990 PREM #108: The 1990 Alternative Questionnaire Experiment: Preliminary Report of the 100-percent Items, December 9, 1991.

Bates, Nancy A. and Theresa J. DeMaio, "Using Cognitive Research Methods to Improve the Design of the Decennial Census Form." Proceedings of the Census Bureau's Fifth Annual Research Conference, 1989, pp. 267-285.

Bates, Nancy A. and Theresa J. DeMaio, 1988 PREM #57: Results of the 1988 National Census Test, January 21, 1992.

DeMaio, Theresa J., "Early Results of the Questionnaire Design Project: Observations of Census Employees Filling Out Census Forms," Internal Census Bureau Memorandum, August 13, 1986.

DeMaio, Theresa J., Elizabeth A. Martin, and Elizabeth Page Sigman, "Improving Response to the Decennial Census Questionnaire." Proceedings of the Section on Survey Research Methods (American Statistical Association), 1988, pp. 161-166.

Martin, Elizabeth, Theresa J. DeMaio, and Pamela C. Campanelli, "Context Effects for Census Measures of Race and Hispanic Origin." Public Opinion Quarterly, Winter 1990, Vol. 54, pp. 551-556.

Table 1. Mail Response Rate by Panel

		Panel				
	Total	1	2	3	4	5	6

Mail Response Rate	50.9	48.2	48.6	49.9	51.8	52.5	54.3
Number of Returned Questionnaires	21160	3340	3369	3460	3589	3645	3757

$X^2=79.6$ d.f.= 5 $p<.001$

Table 2. Components of Nonresponse by Panel

Nonresponse Component	Total	Panel 1	2	3	4	5	6	X^2	d.f.	p value
1) Percent of Forms not Mailed Back	49.1	51.8	51.4	50.1	48.2	47.5	45.7	79.6	5	.001
2) Percent of Returned Forms Missing 100% and Sample Data	2.1	0.5	1.4	1.0	1.8	5.2	2.2	251.9	5	.001
3) Percent of Returned Forms Missing All 100% Person Data (but not sample)	0.4	0.2	1.4	0.7	0.1	0.1	0.0	111.6	5	.001
4) Percent of Returned Forms Missing Sample Person Data (but not 100%)	2.4	4.3	3.9	3.7	1.8	0.6	0.5	228.4	5	.001
5) Percent of Returned Forms Missing all Housing Data	2.8	1.4	1.3	1.4	4.0	3.1	5.3	184.7	5	.001
Total Nonresponse to 100% Person Data (1 + 2 + 3)	51.6	52.5	54.2	51.8	50.1	52.6	47.9	67.8	5	.001
Total Nonresponse to Sample Person Data (1 + 2 + 4)	53.6	56.6	56.7	54.8	51.8	53.3	48.4	139.8	5	.001
Total Nonresponse to Housing Data (1 + 5)	51.9	53.2	52.7	51.5	52.2					

50.6	51.0	14.1	5	.05							
	N				21160	3340	3369	3460	3589	3645	3757

Table 3. Average Percent Item Nonresponse for Various Questionnaire Sections by Panel

				Panel						
			p							
	Total	1	2	3	4	5	6	F-value	value	
100-percent Person Section	3.5	5.8	4.3	3.6	3.3	2.0	2.1	189.8	.001	
N	52837	8313	8356	8664	9128	8851	9525			
Sample Person Section	11.7	12.4	12.5	12.6	12.6	10.1	10.4	35.1	.001	
N	51910	7980	8171	8439	9031	8807	9482			
Housing section - Overall	8.5	8.4	8.1	8.5	10.7	8.3	7.6	16.9	.001	
- Excluding Utility Items	4.9	5.4	4.6	5.0	5.4	4.5	4.3	3.9	.01	
N	20568	3295	3325	3412	3446	3531	3559			

Table 4. Summary of Information Related to the Household Roster by Panel

					Panel					
	Total	1	2	3	4	5	6	X^2	d.f.	p value
No Response to the Household Roster	2.1	5.3	2.2	2.3	2.0	0.3	0.4	280.2	5	.001
N	20568	3295	3325	3412	3446	3531	3559			
Mean Number of Names on Household Roster	2.6	2.5	2.5	2.6	2.6	2.7	2.6	F-value=3.4		.01
Percent of Cases with Eight or More Names on Household Roster	1.0	0.6	1.1	1.1	1.0	1.3	1.3	13.0	5	.05
N	20146	3119	3252	3333	3376	3521	3545			
Age Specific Sex Ratios: Males to Females Ages 15-29	.93	.89	.91	.98	.88	.96	.93			N.S.
N	10727	1594	1628	1698	1831	1904	2072			

Table 5. Summary of Results for the Coverage Items by Panel

				Panel					
	Total	1	2	3	4	5	6	X^2	d.f.	p value
Undercoverage:										
Percent of Forms with One or More Names Added to Undercoverage Question	2.3	0.8	1.9	1.7	2.2	4.9	N.A.	151.9	4	.001
(Number of Names Added)		(29)	(84)	(77)	(97)	(244)				
N	17403	3340	3369	3460	3589	3645				
Overcoverage:										
Percent of Persons Potentially Ineligible	2.8	0.4	4.5	4.6	1.5	2.9	3.1	330.0	5	.001
N	52837	8313	8356	8664	9128	8851	9525			
Distribution of Ineligible "Other Place" Categories										
Another House/Apt.	64.3	--	81.4	82.3	72.2	39.8	36.1	362.2	16	.001
Away at College	23.2	--	13.8	12.7	21.8	39.1	35.8			
In an Institution	1.9	--	1.0	0.5	3.8	3.1	2.7			
At a Military Base	4.1	--	3.8	4.5	2.3	4.3	4.7			
Other	6.6	--	--	--	--	13.7	20.7			

N	1494	31	373	401	133	257	299

REDESIGNING THE CENSUS LONG FORM:

RESULTS FROM THE 1990 ALTERNATIVE QUESTIONNAIRE EXPERIMENT

by

Theresa J. DeMaio and Nancy A. Bates

U. S. Bureau of the Census

Paper prepared for presentation at the
Annual Meetings of the American Statistical Association
Boston, Massachusetts
August 9-13, 1992